Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi
Pamukkale University Journal of Engineering Sciences

Detection of colon cancer using k-means and deep learning algorithms on histopathological images

Histopatolojik görüntüler üzerinde k-ortalamalar ve derin öğrenme algoritmaları kullanılarak kolon kanseri tespiti

Ulaş YURTSEVER[1], Hayrettin EVİRGEN[2], Mustafa Cihat AVUNDUK[3]*

[1]Department of Computer Engineering, Sakarya University, Sakarya, Türkiye.
ulas@sakarya.edu.tr
[2]Faculty of Open and Distance Education, İstanbul University, İstanbul, Türkiye.
hayrettin.evirgen@istanbul.edu.tr
[3]Department of Pathology, Meram Schools of Medicine, Necmettin Erbakan University, Konya, Türkiye.
mcavunduk@hotmail.com

Received/Geliş Tarihi: 24.01.2024 Revision/Düzeltme Tarihi: 16.12.2024 doi: 10.5505/pajes.2025.71508
Accepted/Kabul Tarihi: 07.01.2025 Research Article/Araştırma Makalesi

Abstract

In this research, a novel approach for classifying colon cancer was developed by employing two convolutional neural network (CNN) models, namely GoogLeNet and AlexNet. This approach involves training CNNs with histopathological images segmented into color clusters using an augmented k-means clustering algorithm, rather than utilizing original-raw images. This method was applied to 20 datasets with distinct structural and characteristic features, derived from larger datasets comprising both original and segmented images. The datasets were used to train and test CNN models. The results indicate that AlexNet, trained with segmented images, showed a 2% to 23% increase in accuracy performance, while GoogLeNet's accuracy performance improved by 2% to 27%. Notably, the proposed approach yielded higher accuracy with datasets containing non-homogeneous data.

Keywords: Convolutional neural network, Deep learning, Image segmentation, Image classification, Colon cancer

Öz

Bu araştırmada, GoogLeNet ve AlexNet olmak üzere iki evrişimli sinir ağı (CNN) modeli kullanılarak kolon kanserinin sınıflandırılması için yeni bir yaklaşım geliştirilmiştir. Bu yaklaşımda CNN'ler, orijinal ham görüntüleri kullanmak yerine, artırılmış bir k-ortalamalar kümeleme algoritması kullanılarak renk kümelerine ayrılmış histopatolojik görüntüleri kullanarak eğitilmektedir. Bu yöntem hem orijinal hem de bölütlenmiş görüntülerden oluşan daha büyük veri kümelerinden elde edilen farklı yapısal ve karakteristik özelliklere sahip 20 veri kümesine uygulanmıştır. Veri kümeleri CNN modellerini eğitmek ve test etmek için kullanılmıştır. Sonuçlar, bölümlere ayrılmış görüntülerle eğitilen AlexNet'in doğruluk performansında %2 ile %23 arasında bir artış gösterdiğini, GoogLeNet'in doğruluk performansının ise %2 ile %27 arasında iyileştiğini ortaya koymuştur. Özellikle, önerilen yaklaşım homojen olmayan verilere sahip veri kümelerinde daha yüksek doğruluk sağlamıştır.

Anahtar kelimeler: Evrişimli sinir ağı, Derin öğrenme, Görüntü bölütleme, Görüntü sınıflandırma, Kolon kanseri

1 Introduction

Cancer is the second leading cause of death in Turkey and worldwide. Türkyılmaz et al. [1],[2] note that rising world population, the increase in the numbers of the elderly, significant levels of exposure to potential causes of cancer and various environmental factors contribute to an increase in the number of new cancer cases and are thus expected to aggravate the burden posed by cancer in the future. A review of reported cases of cancer through years clearly shows the increasing burden associated with this condition. Figure 1 presents the data for new cases of six major types of cancer over the years, while Figure 2 shows the number of deaths caused by cancer, with reference to statistics from the GLOBOCAN database, published by the International Agency for Research on Cancer (IARC) under the World Health Organization (WHO) umbrella. A review of these figures to assess the cancer burden reveals that lung, breast and colorectal cancers rank at the top of the list in terms of global incidence rates. Colorectal cancers also rank second in terms of deaths caused by cancer [3],[4],[5],[6]. Thus, it is obvious that early and definite diagnosis of colon cancer is important.

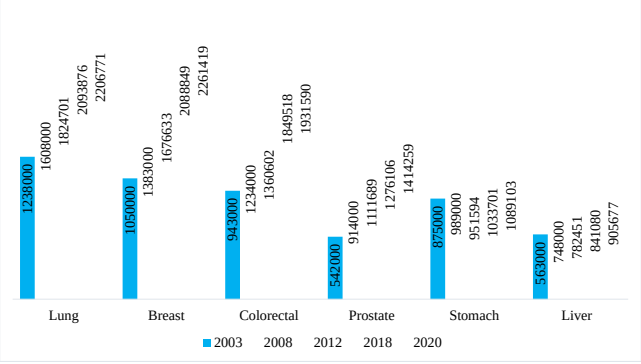

Figure 1. The number of new cancer cases through the years, based on GLOBOCAN database figures [3],[4],[5].

Pathological analysis plays a most crucial part in definite diagnoses of colon tumors. This process entails the analysis of the biopsy samples taken from the relevant tissue, performed

*Corresponding author/Yazışılan Yazar

by a specialized pathologist using a microscope to review the cells and the position of the nuclei, along with structural and functional deformations, culminating in a definite judgment about the nature of the tissue under review.

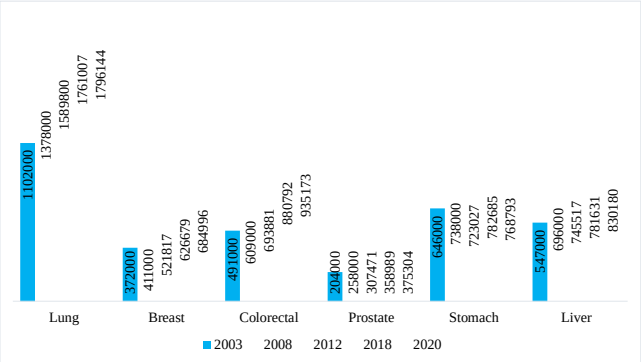

Figure 2. The number of deaths caused by cancer through the years, based on GLOBOCAN database figures [3],[4],[5].

The pathological diagnosis process begins with an assessment of macroscopic and microscopic characteristics. These characteristics are analyzed with reference to the differentiations observed in cells, tissues, and organs [7]. This review is essentially based on visual interpretation and therefore can lead to inconsistent interpretations and non-standard and non-objective results. Variations and a multitude of interpretations are considered common occurrences in efforts to classify the cancer [8]. Therefore, various image analysis systems have been developed to provide objective results, facilitating standardized assessments by physicians. The most important procedures implemented by these image analysis systems, in turn, are image segmentation and image classification [6].

Convolutional neural networks (CNNs) which gained popularity in recent years, and which have been employed in a wide range of applications (face recognition, segmentation, object detection, natural language processing etc.) also see extensive use for image classification as well. However, CNN did not draw much attention in the field of computer vision until 2012, as it required a lot of training data and powerful hardware required to process the significant amount of data involved. However, from 2012 on, given the ability to derive much larger data sets in several distinct fields, the increased emphasis on the big data concept, increased efficiency of algorithms, the development of GPUs, and the fall in computing costs, CNN skyrocketed in popularity [9].

The burgeoning field of digital pathology has seen significant advancements with the integration of deep learning algorithms. Despite these advancements, accurately classifying colon cancer using histopathological images remains a significant challenge due to the complexity and variability of the data. In the literature, the deep learning algorithms working with histopathology images often utilize the original-raw RGB images taken in histopathology analyses as input. While alternative color spaces have been explored in some studies, raw RGB images remain the most common input for deep learning models.

Several studies have explored the potential of deep learning architectures, particularly convolutional neural networks (CNNs), for colon cancer classification. For example, Parelanickal et al. [10] achieved 95% and 97% accuracies using VGG16 and MobileNetV2, respectively. Sari et al. [11] demonstrated effective results comparing CNNs with Vision Transformers, illustrating modern architectures' potential. Peng and Lee [12] achieved 99.77% accuracy with ResNet50 in histopathological image analysis, showcasing pre-trained networks' efficacy. Anju and Vimala [13] reported a 97.42% accuracy for InceptionV3, highlighting the impact of preprocessing and data augmentation on performance. Collectively, these studies underscore deep learning architectures' significant potential in colon cancer classification.

Moreover, fine-tuning models like AlexNet and GoogLeNet enhances classification accuracy for colon cancer datasets. Optimized models show notable performance improvements in accuracy, sensitivity, and specificity. Babu et al. [14] found that data augmentation increased GoogLeNet's accuracy by 2.3% to 80% on the CrchistophenoTypes dataset, demonstrating its role in enhancing generalization.

Kumar et al. [15] emphasized preprocessing steps, showing that median filtering and SegNet-based preprocessing improved AlexNet and GoogLeNet's classification accuracy. Their comparative evaluation revealed GoogLeNet as the best performer with 94.165% accuracy, 97.589% sensitivity, and 87.359% specificity, illustrating its high sensitivity in colon cancer classification when fine-tuned.

Conversely, Kumar et al. [16] showed that optimizing AlexNet with the WSO method raised its classification accuracy to 99.3%, with sensitivity and specificity values of 99.76% and 98.68%. This indicates that while GoogLeNet is effective, AlexNet can achieve superior accuracy through specific optimization strategies.

In summary, CNN architectures like AlexNet and GoogLeNet are prevalent in colon cancer classification. However, optimizing these models and tailoring them to datasets are crucial for enhancing accuracy and generalization performance. The literature supports the efficacy of data augmentation and preprocessing techniques in this field.

Nonetheless, current research primarily addresses raw, unsegmented images. While prior research predominantly utilizes unsegmented images, some studies explore alternative color spaces (HSV, YCbCr and RGB [17]; RGB, YCbCr, CIELab and HSV [18]; CIELab* and RGB [19]; RGB and HIS [20]). Research indicates that image conversion can enhance CNN classification accuracy. However, the extent of enhancement is often contingent upon the model architecture employed.

Although these studies demonstrate the potential of CNNs and alternative preprocessing techniques, they largely focus on raw or color-space-modified images. To address this gap, the proposed method enhances model performance by segmenting RGB images using an augmented k-means clustering algorithm [21], thus reducing color complexity levels without altering the color space. Moreover, the tests performed found performance improvements with various CNN models. Moreover, even with problematic datasets impairing the performance of the model and causing overfitting in the system, the use of the method proposed here was seen to bring about 27% improvement in system performance.

The structure of this paper is outlined as follows: Section 2 details the dataset, including image acquisition, the characteristics of the images obtained, preprocessing, and augmentation methods. Subsection 2.4 elucidates the augmentation k-means clustering algorithm for segmentation, while Subsection 2.5 presents the CNN architectures utilized

for classification. Section 3 offers a comparative analysis of CNN models with segmented and raw images. Finally, Section 4 summarizes the findings and emphasizes the principal findings.

2 Materials and methods

The methodology consists of two main phases: image preprocessing and CNN training. In the preprocessing stage, an augmented k-means clustering algorithm is applied for image segmentation. This algorithm simplifies the color scheme of each image, reducing color complexity and highlighting important features. After segmentation, these images are used to train two CNN models: GoogLeNet and AlexNet. The training process is conducted with a diversified dataset, ensuring a comprehensive learning process.

2.1 Microscope image

The study is based on color images of human colon tissues, taken with a digital camera (Nikon Coolpix E5000) mounted on a microscope (Nikon Eclipse E400). All images used in the study are microscope images processed with the H&E [22],[23] staining technique. H&E staining is the most popular staining technique in histopathology. H&E staining dyes the histopathology images with several dyes affecting specific parts of the tissue: blue and black showing the cell nuclei, pink presenting the cytoplasm, dark red in the case of muscle fibers, and an orange-red hue showing the red blood cells (RBCs).

A study on the segmentation of microscope images using augmented k-means clustering algorithm [6] used microscope images of the colon tissue, with a size of 2560x1920 pixels. 51 of these images belonged to benign colon tumors, and 49 belonged to malignant colon tumors. The microscope images are a random sample of the dyed slides of the tissues already diagnosed and archived by the Medical Pathology Department of Necmettin Erbakan University. Furthermore, the approval (document ID 2015/116 and 13 Feb 2015 dated) of Necmettin Erbakan University Ethics Committee for Non-Pharmaceutical and Non-Medical Device Research was obtained for the use of these images.

The classification of tumors as malignant and benign tumors using deep learning with segmented images, on the other hand, the dataset was increased to a total of 123 images (62 benign and 61 malignant tumor images) with additions alongside the images used in the segmentation analysis. Moreover, images provided on the website of the Department of Computer Science at the University of Warwick (https://warwick.ac.uk/fac/sci/dcs/research/tia/glascontest/) covered by a consent for use in research were used. These additional images comprise a total of 165 shots obtained from 16 histology slides with H&E staining, taken from different patients. To digitize the histology slides into a whole-slide image form, Zeiss Mirax Midi Slide Scanner with a pixel size of 0.465 µm pixel. The whole-slide images thus digitized are scaled to a pixel size of 0.620 µm, equivalent to 20x lens magnification [24],[25]. 74 of these images belonged to benign colon tumors, and 91 belonged to malignant colon tumors.

2.2 Computer hardware and software

Segmentation or classification of an image at a high level of detail poses a very complex problem. In the same vein, the use of deep learning algorithms for the classification of images with a high level of detail also represents a highly complex problem requiring significant processing power. Powerful hardware capable of meeting significant processing requirements is used. Graphics processing units (GPUs) are employed, given their architecture based on very high numbers of parallel processors. Thus, GPUs stand out as crucial hardware for deep learning algorithms. In the present study, the deep learning system working with segmented images was a PC with Gigabyte X299 Aorus Gaming 9 main board, 3.5 Ghz Intel Core i7 7800X processor for socket 2066, with 8.25MB cache, Kingston DDR4 64GB RAM 3000Mhz, 2 x GeForce GTX 1080Ti Gaming X-Trio GDDR5X 11GB 352 Bit Nvidia graphics cards installed in an SLI arrangement, and 3 x 4TB Sata 3.0 128MB cache 7200 Rpm NAS hard drives. The software used was NVIDIA Digits 6 [26], a web-based software running on Ubuntu 18.10 operating system. The parallel processing architecture was CUDA 9.0 [27], which increased processing performance by using NVIDIA GPU, complemented with the CuDNN [28] library enabling the tasks required for deep learning. The software was installed on the Docker, which is the recommended solution for virtualization at the operating system level. Moreover, Caffe v0.15.14 [29],[30] based on Python programming language served as the deep learning library, while Stochastic Gradient Descent (SGD) algorithm, which is the most popular choice in optimization processes, was used for determining the minimum level in the loss function.

2.3 Data augmentation and pre-processing

Data augmentation refers to the process whereby a smaller dataset is used to produce a larger one, all the while maintaining the ability to represent real-world data. Several data augmentation methods have been proposed for this purpose. Many images are required in architectures where CNN, a leading DL algorithm, is used. If the dataset contains only a limited number of images, data augmentation could help against overfitting.

In the present study, with a view to increasing the number of images, first of all, the 62 images of benign colon tumors and 61 images of malignant colon tumor, each with a size of 2560x1920 pixels, received from the archives of the Medical Pathology Department of Necmettin Erbakan University, were processed to produce completely distinct 464 images of benign tumors and 407 images of malignant tumors, based on different sections of the original-raw images. In the same vein, 74 images of benign colon tumors and 91 images of malignant colon tumors published on the website of the Department of Computer Science at the University of Warwick and covered by a consent for use in research were used to produce completely distinct 354 images of benign tumors and 424 images of malignant tumors, based on different sections of the original-raw images. This process resulted in a dataset of 818 images of benign colon tumors, and 831 images of malignant colon tumors.

These microscopic images of colon tumors were then subjected to data augmentation in a process imitating the cases with real world data, through vertical and horizontal flipping, 90^{o} 180^{o} and 270^{o} rotation, and adding gaussian noise. Figure 3 presents examples of images obtained through data augmentation on a malignant tumor image. By applying the data augmentation methods shown in Figure 3, the number of images in the dataset was increased to a total of 11543, comprising 5726 benign and 5817 malignant tumor images.

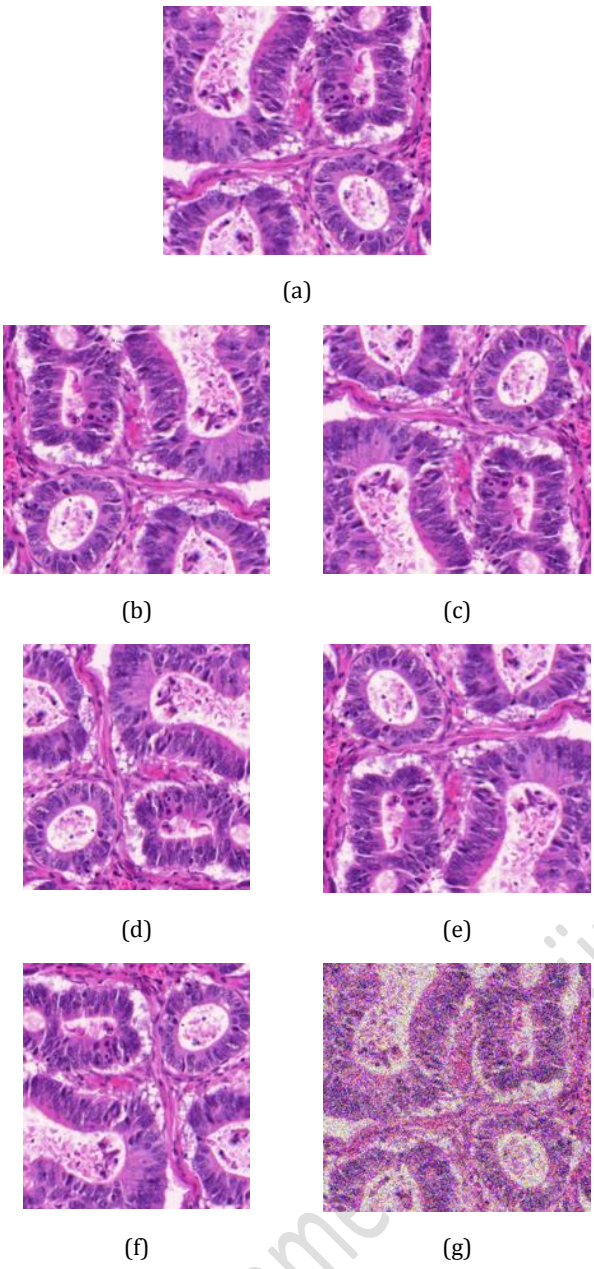

Figure 3. Data augmentation applied on a malignant tumor image: (a) original image, (b) horizontal flip, (c) vertical flip, (d) 90° rotation, (e) 180° rotation, (f) 270° rotation, (g) gaussian noise.

2.4 Augmented k-means clustering algorithm

Determining the number of clusters involved and defining the pixel clusters is a difficult and important step in the segmentation of color images through clustering-based processes. In particular, accurate and optimal determination of the number of clusters and the initial cluster centers is crucial in terms of performance, segmentation quality, and segmentation success. Moreover, from a system stability perspective, it is imperative to come up with identical results when producing images through segmentation using clustering methods, in other words, to get the same clusters and the same pixel values associated with those clusters in every reiteration of the process. In medicine getting accurate and consistent results in every iteration using an identical histopathology image is crucial regarding the system's reliability. The initial cluster centers should be chosen based on a specific and well-defined algorithm to ensure such consistency. Yet, when doing so, one should also consider the need to avoid excessive processing requirements and present a simple and stable structure. For this purpose, to ensure an optimal distance and separation between the initial cluster centers, the augmented k-means clustering algorithm proposed in the previous study was used [6]. This algorithm is essentially a further development on the weighted k-means clustering algorithm [31],[32]. Processing steps of the augmented k-means clustering algorithm:

Step 1: First, the number of clusters (k) to group the data elements (pixels in the case of images) into will be established.

Step 2: The initial cluster center (C_k) (centroid) value will be established for each cluster, based on Equation 1. Doing so ensures that the initial cluster centers are located at an optimal distance from each other.

$$K_c = \frac{D_h}{k+1} \qquad (1)$$

$$for\ \ i = (1, \dots, k), \qquad C_i = K_c * i$$

Here, K_c refers to the cluster center creation coefficient, D_h refers to the number of bins in histogram, k refers to the cluster number and C_i refers to the cluster center value.

Step 3: The distance between all data elements and the cluster centers will be calculated with the Euclidean distance formula provided in Equation 2. All data elements will then be assigned to the nearest cluster center, based on the results thus produced.

$$d(a,b) = \sqrt{\sum_{i=1}^{3}(a_i - b_i)^2} \qquad (2)$$

When calculating the distance between the two colors (a, b), the i value in Equation 2 stands for the color channel, as the color values of the color images in the dataset are expressed as RGB values.

Step 4: The cluster center values for these sets which are now populated with new elements will then be recalculated using the formula presented in Equation 3. The new cluster centers thus found will be compared against the previous cluster center. In this equation, the values of the elements of the set, along with the histogram data for those elements, are also included in the formula.

$$C_k = \frac{1}{\sum_{i=1}^{m} h_i} \sum_{j=1}^{m} h_j Z_j \qquad (3)$$

Here, k refers to the cluster number, C = {C_1, C_2, ..., C_k} refer to the cluster center values, m refers to the number of elements assigned to each cluster center, $C_{\{1\dots k\}}$ = {Z_1, Z_2, ..., Z_m} refer to the elements corresponding to the cluster center, and h_j refers to the histogram value for the color Z_j.

Step 5: In case the cluster centers are the same with the previous cluster center values, the clusters will be assumed to have remained unchanged, and the process shall conclude. In case any changes are observed with the cluster centers, all steps from step 3 on will be repeated.

2.5 Convolutional neural network

Convolutional neural network (CNN) is a deep learning algorithm, which is a type of artificial neural networks (ANN) and refers to a specific type of deep feedforward neural networks, characterized by a grid-like structure and often used for data processing in the context of analyzing images. In other words, it refers to neural networks employing convolution as a mathematical operation, instead of matrix multiplication in layers [33].

Deep learning offers solutions based on deep neural network algorithms for complex problems requiring variable functions in voice recognition, face recognition, visual object recognition, object detection, bioinformatics, natural language processing, image segmentation, and classification, as well as for segmentation, detection, classification, and registration problems regarding medical imaging [9],[34]-[37]. As one such deep neural network algorithms, CNN finds extensive use in several fields such as classification, object detection, and natural language processing, due to its high performance in computer vision. The first layers of CNNs often derive simpler and cruder properties. However, as the number of layers increase, a range of more complex properties can also be derived in deeper layers. Such complex properties are the ones representing the image much better [38].

Yet, CNN did not draw much attention in the field of computer vision till 2012, as it required a lot of training data and powerful hardware required to process the significant amount of data involved. In 2012, after Krizhevsky et al. won the top place in ImageNet image classification competition, modern deep learning models gained popularity and began to be used in various fields [9]. LeNet [39] and AlexNet [40] were the first modern deep learning models introduced in this paradigm and were soon followed by several others including ZFNet [41], GoogLeNet [42],[43], VGGNet [44], and ResNet [45]. AlexNet and GoogLeNet models were used extensively and can be considered popular models. The present study is based on AlexNet and GoogLeNet models.

In this study, AlexNet and GoogLeNet architecture were preferred due to their widespread applications in the literature and compatibility with existing resources The modularly designed architecture of GoogLeNet is a significant advantage over complex and more resource-intensive models such as ResNet in terms of reduced computational complexity and architectural flexibility. This attribute renders it especially beneficial in scenarios where computational resources are constrained or where the minimization of computational expenses is of paramount importance. Nevertheless, research has demonstrated that GoogLeNet exhibits competitive performance across a spectrum of tasks while sustaining a judicious equilibrium between accuracy and computational efficiency [42,43].

In a similar vein, AlexNet [40], recognized as one of the foundational architectures in deep learning, is often favoured its status as a benchmark model within the domain of image classification. Its comparatively straightforward architecture renders it an exemplary option for conducting comparative analyses. The employment of AlexNet offers a robust framework for the validation of the proposed methodologies by facilitating a direct comparison of the results of this study with existing research.

While more contemporary architectures, such as ResNet, are engineered to attain superior accuracy on more complex tasks, they frequently necessitate greater computational resources and larger datasets. Therefore, in this study, where the primary aim was to evaluate the impact of using histopathological images segmented into color clusters by an augmented k-means clustering algorithm instead of original raw images on the performance of CNNs, AlexNet and GoogLeNet architectures were preferred over more advanced models such as ResNet. This preference was based on their ability to provide an optimal balance between computational efficiency and classification accuracy. For these reasons, these architectures were deemed suitable for achieving the objectives of this study.

The AlexNet and GoogLeNet architectures were trained on the dataset using the Nvidia-DIGITS 6 platform, which applied default hyperparameter settings. The learning rate was set to 0.01, batch size to 32, and the number of epochs to 30. The models were initialized with standard random weights since no pre-trained networks were employed.

2.5.1 AlexNet

Developed by Alex Krizhevsky, Ilya Sutskever and Geoff Hinton, the AlexNet model [40] won the ImageNet challenge in image classification in year 2012, with an error rate of 16.4%, and is based on multiple convolutions. It has an 8-layer architecture, 5 of which are convolution layers, complemented by 3 fully connected layers. The activation function employed is Rectified Linear Unit (RELU). Figure 4 presents the architecture of AlexNet.

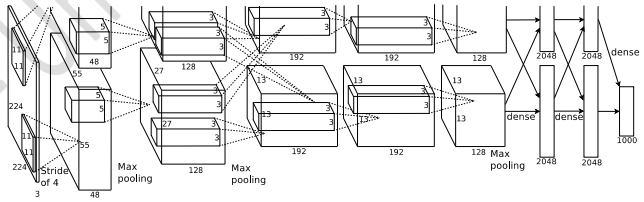

Figure 4. AlexNet architecture [40].

The initial convolutional layer processes the 224×224×3 input image utilizing 96 filters with dimensions of 11×11×3 and a stride of 4, culminating in an output dimension of 55×55×96. Subsequent convolutional layers systematically enhance the feature maps by employing diminutive filter sizes (5×5 or 3×3) and diverse strides. The second convolutional layer utilizes 256 filters measuring 5×5×48 on the normalized and pooled output generated by the first layer. The third convolutional layer implements 384 filters of size 3×3×256 on the output produced by the second layer. The fourth convolutional layer employs 384 filters of size 3×3×192, while the fifth convolutional layer utilizes 256 filters of size 3×3×192. In the aftermath of the convolutional layers, the architecture incorporates three fully connected layers, each comprising 4096 neurons, facilitating feature extraction and classification [40].

2.5.2 GoogLeNet

GoogLeNet [42],[43] architecture was developed by Google employees and introduced as a participant in the ImageNet challenge in 2014. GoogLeNet architecture has 22 layers and can operate with 12 times fewer parameters than AlexNet while achieving a higher accuracy (with an error rate of 6.7%). The basic building block of this architecture is the inception module. The architecture contains a total of 9 inception modules. The block diagram showing the inception module of the GoogLeNet architecture is presented in Figure 5.

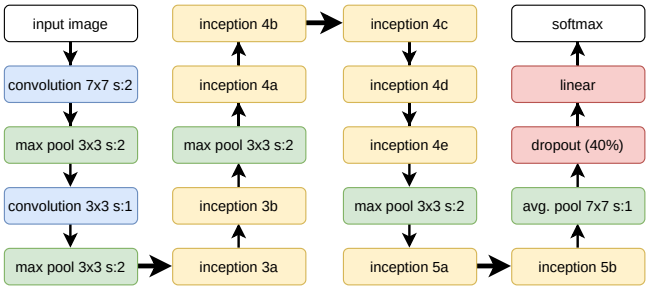

Figure 5. GoogLeNet block diagram.

In Figure 5, each number corresponds to the size of the feature maps, filter dimensions, or stride values used in the respective layers. This modular and hierarchical architecture facilitates GoogLeNet in attaining superior accuracy while utilizing a reduced number of parameters, thereby rendering it exceptionally effective for extensive image classification endeavors.

The input image adheres to the standard RGB format and measures 224×224×3. The initial convolutional layer employs a 7×7 filter with a stride of 2, diminishing the input's spatial dimensions. Subsequent convolutional layers utilize 3×3 filters to capture features at more granular scales. Max-pooling layers follow the initial convolutional layers to condense spatial dimensions while preserving critical features. For instance, the first max-pooling layer decreases the dimensions to 56×56×64.

The inception modules, which form the core of GoogLeNet, are represented in yellow boxes (e.g., Inception 3a, 4b). These modules execute concurrent convolutions with filters of sizes 1×1, 3×3, and 5×5, alongside max-pooling. The resultant outputs from these processes are amalgamated to constitute the ultimate output of each inception module. For instance, Inception 3a combines outputs from multiple filter sizes to create a feature map that captures information at various scales.

To improve training and prevent vanishing gradients, auxiliary classifiers (e.g., Inception 4a and 4d) are included. These classifiers provide intermediate supervision during training. After passing through the inception modules, the network transitions to fully connected layers, incorporating a dropout rate of 40% to reduce overfitting.

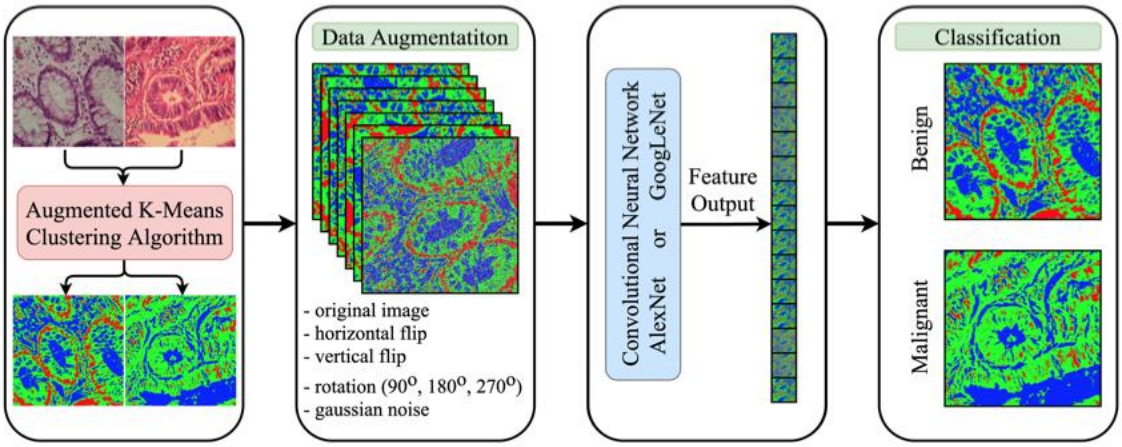

Figure 6. Overall scheme of the system

3 Results and discussion

CNNs, a kind of deep learning algorithms gaining popularity in recent years allow the feeding of the images in their original-raw forms as input for the neural network model, without any pre-processing other than data augmentation, for the purposes of image classification in general, and histopathology image classification in particular. As seen in Figure 6, the present study proposed the use of images after some processing to reduce their detail levels prior to the classification process, rather than using the full original image with all the details it offers, on the grounds of simplifying the procedures and improving the performance in the context of the efforts to classify histopathology images as images of malignant or benign colon tumors. To do so, a fast and simple method of segmentation process was applied to reduce the level of details all the while processing the basic properties of the original-raw images. The segmentation was based on a proposed augmented k-means clustering algorithm to operate in a simple, fast, and consistent manner.

The original-raw images of benign and malignant tumor tissues were subjected to augmented k-means clustering algorithm to produce 4 distinct models in terms of cluster count, with k = {3, 4, 5 and 6} for each of the respective iterations. In the segmentation process, the colors assigned to the clusters in accordance with the cluster count were **red**, **green,** and **blue** for the 3-cluster model; **red**, **yellow**, **green,** and **blue** for the 4-cluster model; **red**, **yellow**, **green**, **dark green,** and **blue** for the 5-cluster model; and **red**, **yellow**, **green**, **dark green**, **blue,** and **light blue** for the 6-cluster model.

Then came a comparative analysis of the impact on the model performance, of the use of un-preprocessed original-raw colon cancer microscope images, against the use of RGB (3 channel) images segmented through the present algorithm on CNNs. The comparative analysis was performed on a deep learning algorithm, namely a CNN, based on AlexNet and GoogLeNet architectures.

First, through random selection, 70% of the dataset comprised of the original-raw images were assigned to the training set, 15% to the validation set, and 15% to the test set. In the next step, the AlexNet and GoogLeNet models were trained using the dataset. The process was repeated through 10 iterations. In each iteration the training, validation and test datasets were refreshed, and the models were trained on the new datasets

thus presented. Through the iterations, the low levels of model performance associated with specific datasets were noted. In the present study, the dataset noted for a particularly low level of performance was the "ozel" dataset. Distinct sub-images were produced by segmentation through the cropping procedure applied on the original-raw images. The dataset comprised of all sub-images derived from a given original-raw image is designated as the "bir" dataset, whereas the dataset comprised of sub-images derived from the images distributed randomly, without the requirement to contain all such sub-images associated with a given original-raw image is designated as the "par" dataset.

Based on these arrangements, the original-raw images numbering 11543 in total, and the segments thereof through the use of 3, 4, 5, and 6 color clusters were used to produce 20 image clusters with distinct structures and characteristics, named bir, bir_ozel, par, and par_ozel based on each image cluster. These datasets are designated as orj_bir, orj_bir_ozel, orj_par, orj_par_ozel, k3_bir, k3_bir_ozel, k3_par, k3_par_ozel, k4_bir, k4_bir_ozel, k4_par, k4_par_ozel, k5_bir, k5_bir_ozel, k5_par, k5_par_ozel, k6_bir, k6_bir_ozel, k6_par and k6_par_ozel. Following their creation, these datasets were then divided into three categories: the training dataset, the validation dataset, and the test dataset. All the generated datasets were subsequently utilized to train and test the AlexNet and GoogLeNet CNN models.

Table 1. The confusion matrix obtained from testing AlexNet and GoogLeNet on the test dataset.

	AlexNet								GoogLeNet							
Dataset	TP	TN	FN	FP	b (%)	m (%)	Abs (%)	Acc (%)	TP	TN	FN	FP	b (%)	m (%)	Abs (%)	Acc (%)
orj_par dataset	830	24	25	843	97.19	97.14	0.05	97.15	825	29	22	846	96.6	97.47	0.87	97.04
orj_par_ozel dataset	386	468	13	855	45.2	98.5	53.3	72.07	372	482	16	852	43.56	98.16	54.6	71.08
orj_bir dataset	898	54	85	727	94.33	89.53	4.8	92.12	872	80	51	761	91.6	93.72	2.12	92.57
orj_bir_ozel dataset	357	539	34	771	38.84	95.78	56.94	66.31	379	517	45	760	42.3	94.41	52.11	66.96
Avg. acc. for original-raw images					68.89	95.24	28.77	81.91					68.52	95.94	27.43	81.91
k3_par dataset	790	64	49	819	92.51	94.35	1.84	93.44	824	30	20	848	96.49	97.7	1.21	97.1
k3_par_ozel dataset	795	59	32	836	93.09	96.31	3.22	**94.72**	765	89	20	848	89.58	97.7	8.12	**93.67**
k3_bir dataset	902	50	122	690	94.75	84.98	9.77	90.25	922	30	55	757	96.85	93.23	3.62	**95.18**
k3_bir_ozel dataset	734	162	7	798	81.92	99.13	17.21	**90.06**	816	80	18	787	91.07	97.76	6.69	**94.24**
Avg. acc. for 3 cluster segmented images					90.57	93.69	8.01	92.12					93.50	96.60	4.91	95.05
k4_par dataset	812	42	49	819	95.08	94.35	0.73	94.72	827	27	52	816	96.84	94.01	2.83	95.41
k4_par_ozel dataset	697	157	33	835	81.62	96.2	14.58	88.97	794	60	53	815	92.97	93.99	1.02	93.44
k4_bir dataset	871	81	69	743	91.49	91.5	0.01	91.5	903	49	50	762	94.85	93.84	1.01	94.39
k4_bir_ozel dataset	666	230	86	719	74.33	89.32	14.99	81.42	813	83	31	774	90.74	96.15	5.41	93.3
Avg. acc. for 4 cluster segmented images					85.63	92.84	7.58	89.15					93.85	94.50	2.57	94.14
k5_par dataset	835	19	20	848	97.78	97.7	0.08	**97.74**	837	17	24	844	98.01	97.24	0.77	**97.63**
k5_par_ozel dataset	653	201	13	855	76.46	98.5	22.04	87.57	717	137	27	841	83.96	96.89	12.93	90.48
k5_bir dataset	921	31	71	741	96.74	91.26	5.48	**94.22**	919	33	68	744	96.53	91.63	4.9	94.27
k5_bir_ozel dataset	581	315	10	795	64.84	98.76	33.92	80.89	800	96	15	790	89.29	98.14	8.85	93.47
Avg. acc. for 5 cluster segmented images					83.96	96.56	15.38	90.11					91.95	95.98	6.86	93.96
k6_par dataset	798	56	46	822	93.44	94.7	1.26	94.08	838	16	36	832	98.13	95.85	2.28	96.98
k6_par_ozel dataset	682	172	35	833	79.86	95.97	16.11	87.98	735	119	59	809	86.07	93.2	7.13	89.66
k6_bir dataset	907	45	127	685	95.27	84.36	10.91	90.25	912	40	86	726	95.8	89.41	6.39	92.86
k6_bir_ozel dataset	638	258	27	778	71.21	96.65	25.44	83.25	811	85	22	783	90.51	97.27	6.76	93.71
Avg. acc. for 6 cluster segmented images					84.95	92.92	13.43	88.89					92.63	93.93	5.64	93.30

The training based on the AlexNet and GoogLeNet models produced the confusion matrix results presented in Table 1. The model was run 3 times for each dataset, and the training model performing best was employed ultimately. In Table 1, benign tumors are labeled with the letter b, malignant tumors are labeled with the letter m, while accuracy in the diagnosis of benign tumors is b (%), and the accuracy in the diagnosis of malignant tumors is m (%). The absolute margin between the accuracy rates for the benign tumors and malignant tumors is Abs (%). Finally, the model's overall accuracy is shown as Acc (%). The cases where the microscope images of benign tumors were correctly classified as benign were labeled true positive (TP), whereas the cases where they were misdiagnosed as malignant were labeled false negative (FN). In line with the same naming convention, the cases where the microscope images of malignant tumors were correctly classified as malignant were labeled true negative (TN), while the cases where they were misclassified as benign were labeled false positive (FP).

The review of Table 1 reveals that, with "par" and "bir" datasets the highest model accuracy for AlexNet is achieved with the use of segmented images (k=5). In the same vein, with "par_ozel" and "bir_ozel" datasets, the highest model accuracy for AlexNet is achieved with the use of segmented images (k=3) for training the model. In particular, in the case of the datasets named "par_ozel" and "bir_ozel", the AlexNet model trained on segmented images achieved vastly improved model accuracy compared to the AlexNet model trained on original-raw images, with 22% better performance observed with the "par_ozel" dataset, and 23% better performance observed with the "bir_ozel" dataset. Thus, model performance is clearly increased with the use of segmented images for model training and testing processes for the AlexNet model.

The training based on the GoogLeNet model produced the confusion matrix results presented in Table 1. The review of the average accuracy rates of the model as presented in Table 1 reveals that the use of segmented images in general, and images

segmented into three clusters (k=3) in particular, for model training and testing processes leads to increased model performance. When trained on "par" dataset alone, the GoogLeNet model trained with images segmented into 5- and 3-color clusters achieved an accuracy rate less than 1% better than the accuracy of the model trained with original-raw images. Such a small margin is arguably a consequence of the information leak from the training dataset to the test dataset, occurring both with the AlexNet and GoogLeNet models. In the case of the datasets named "par_ozel", "bir", and "bir_ozel", the GoogLeNet model trained on segmented images achieved improved model accuracy compared to the GoogLeNet model trained on original-raw images, with more than 2% better performance observed with the "bir" dataset, and 22% and 27% better performance respectively observed with the "par_ozel" and "bir_ozel" datasets.

The review of the model accuracy rates presented in Table 1, obtained through the testing of the neural network built with AlexNet and GoogLeNet CNN models reveals that the GoogLeNet model performs better in the classification of histopathology images of colon cancer as benign or malignant type cancer. Furthermore, the GoogLeNet model trained on images segmented into three clusters performs better in classification compared to GoogLeNet models trained on other datasets.

Although the confusion matrix results presented in Table 1 suggest that the use of segmented images increases performance of CNN models, it is still imperative to investigate the diagnostic ability associated with such performance levels, using receiver operating characteristic (ROC) analysis used extensively in the field of medicine.

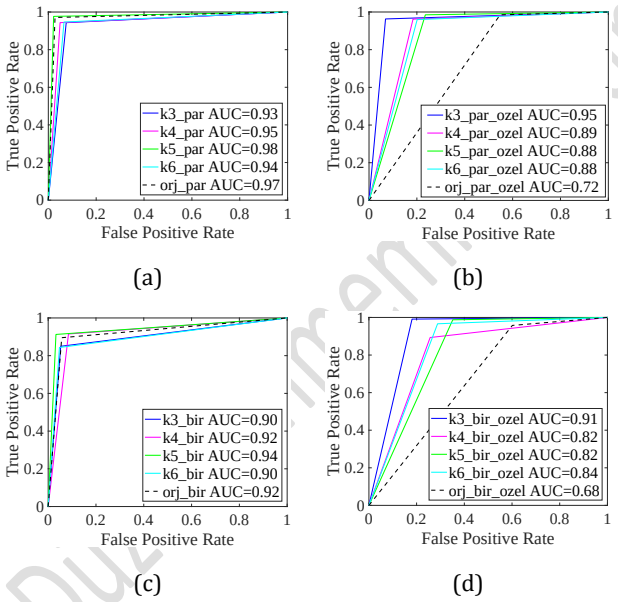

Figure 7. ROC curves of the AlexNet model trained on original-raw images and images segmented with 3, 4, 5, and 6 clusters: (a) "par" dataset, (b) "par_ozel" dataset, (c) "bir" dataset, (d) "bir_ozel" dataset.

ROC analysis is a graphical analysis method used frequently in identifying the model with the best classification performance among various machine learning applications in the field of medicine. In this context, the performance levels of the CNN models ran with distinct datasets labeled "par", "par_ozel", "bir" and "bir_ozel" in the present study were also evaluated using ROC analysis. The ROC analysis assessed the diagnosis performance of the model in question, with reference to the area under the curve (AUC) value.

The ROC curves and AUC values pertaining to the performance results obtained through training of the AlexNet with 20 distinct datasets are presented in Figure 7. A glance at the ROC curves (see Figure 7) showing the diagnosis ability of the AlexNet model reveals that the images included in the "par_ozel" and "bir_ozel" datasets, which cannot be considered good representatives of the sample space for the purposes of classification of colon cancer images, lead to very low AUC values in case they are used without any pre-processing. On the other hand, the training of the AlexNet model using the segmented versions of the same images as per the method proposed in the study, leads to an AUC value approaching 1. In the light of these observations, the ROC graphs run in parallel to the model accuracy rates presented in Table 1.

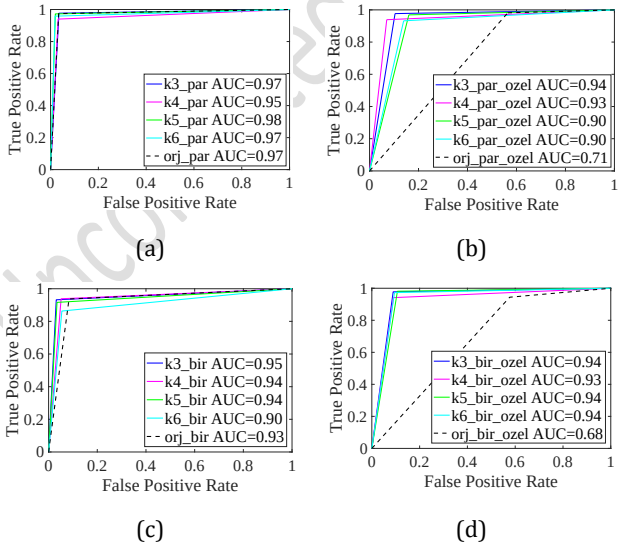

Figure 8. ROC curves of the GoogLeNet model trained on original-raw images and images segmented with 3, 4, 5, and 6 clusters: (a) "par" dataset, (b) "par_ozel" dataset, (c) "bir" dataset, (d) "bir_ozel" dataset.

The ROC curves and AUC values pertaining to the performance results obtained through training of the GoogLeNet with 20 distinct datasets are presented in Figure 8. A glance at the ROC curves (see Figure 8) for the GoogLeNet model reveals that the images included in the "par_ozel" and "bir_ozel" datasets, which cannot be considered good representatives of the sample space for the purposes of classification of colon cancer images, lead to very low AUC values in case they are used without any pre-processing, as was the case with the AlexNet model. Thus, the performance margin between the AlexNet and GoogLeNet models using original-raw versions of the images and the AlexNet and GoogLeNet models using the dataset comprises of images segmented through the method proposed is statistically significant.

The impact of the images used in the datasets on model performance is assessed through the confusion matrix, whereas the ROC curves helped analyze the confidence levels of the results obtained, and to ascertain the optimal match of the CNN models with specific datasets. In the end, it was found that the GoogLeNet model trained with images segmented through augmented k-means clustering performed better. On the other hand, results or errors affecting performance of the successful

model, due to certain distribution patterns of the images across the training and test databases, may have occurred. To eliminate such errors, and to verify the accuracy and reliability of the model performance assessment independent of the specific dataset distribution involved, the data sample was split into k groups, using the k-fold cross validation method. Thereafter, k distinct instances of training and test datasets were produced, with one item being assigned to the test set, and the rest being assigned to the training set. The CNN model was trained separately with each training set, and thereafter tested with the associated test sets. The mean performance level thus assessed represents the true performance of the model. The GoogLeNet model was trained separately with images segmented into 3 clusters, as well as with the original-raw images. With the k-fold cross validation, the k was set to 10 for "par" and "bir" datasets, to 6 for "par_ozel" dataset, and 7 for "bir_ozel" dataset. Doing so allowed the verification of the model's accuracy, independent of the positive or negative potential impact the data in the chosen dataset may have on the model's performance. The average performance figures established using k-fold cross validation for the GoogLeNet model with "par", "par_ozel", "bir", and "bir_ozel" datasets comprised of the original images are presented in Table 2.

Table 2. The comparison of the average performance levels of the GoogLeNet model with original-raw images and images segmented into 3 clusters.

	Segmented image			Original-raw image		
Dataset	b(%)	m(%)	Acc(%)	b(%)	m(%)	Acc(%)
par	97.57	97.72	97.64	96.37	98.02	97.20
par_ozel	97.67	97.49	97.59	87.95	97.94	92.99
bir	95.40	93.75	94.54	90.27	95.84	93.06
bir_ozel	92.89	93.29	93.08	81.86	94.06	88.05
Average	**95.88**	**95.56**	**95.71**	**89.11**	**96.47**	**92.83**

4 Conclusion

The present study is based on color images of human colon tissues, which underwent histology analysis with H&E routine staining. The first step had been to apply a simple and fast segmentation of the images of colon tumors. To do so, augmented k-means clustering algorithm proposed and subjected to a performance assessment in our previous study [6] was employed. After that, a DL-based approach utilizing these segmented images was applied to classify colon tumor images as malignant or benign. Then, the performance of the DL-based classification using segmented versions of histopathology images of colon tumors was compared against the DL-based classification using the original-raw images.

Two CNN models, AlexNet and GoogLeNet, were employed, leading to an innovative method for classifying colon tumors. Within this framework, 136 images of benign tumors and 152 images of malignant tumors, totaling 288 tumor images, were used to generate 818 segmented images of benign tumors and 831 segmented images of malignant tumors, by taking various cross-sections of these images. These processes were followed by data augmentation involving vertical flipping, horizontal flipping, 90^o, 180^o and 270^o rotation, and adding gaussian noise to come up with a sample of 11543 images to represent real world data better for the deep learning processes. To train the deep learning-based CNNs, segmented images divided into color clusters via the segmentation method used in the initial stage were utilized, rather than the original, unprocessed images. To assess the performance of this proposed approach, a total of 20 datasets with distinct structures and characteristics, named after the "par", "par_ozel", "bir" and "bir_ozel" pattern, comprised of original-raw images and segmented images were employed. The datasets created were subsequently employed with the AlexNet and GoogLeNet CNN models for training and testing. These test outcomes were the basis for constructing confusion matrices, plotting ROC curves, and computing AUC values.

The comparison revealed that, with the "par" dataset, the AlexNet model trained with segmented images (97.74%) achieved 0.59% higher accuracy compared to the AlexNet model trained with original-raw images (97.15%). The closeness of the accuracy levels in the two distinct training settings is probably due to information leaks involving the training dataset and the test dataset. With the "bir" dataset, on the other hand, the AlexNet model trained with segmented images was observed to achieve 2.1% better than the model trained with original-raw images. With the "par_ozel" dataset, the AlexNet model trained with segmented images achieved 22.65% higher accuracy compared to the AlexNet model trained with original-raw images, whereas with the "bir_ozel" dataset, the model trained with segmented images achieved 23.75% higher accuracy. Thus, one can conclude that the AlexNet model trained with segmented images achieves up to 23% higher performance.

In case the same datasets are used to train the GoogLeNet deep learning model, the GoogLeNet model trained on segmented images performed less than 1% better accuracy compared to the GoogLeNet model trained on original-raw images. Moreover, compared against the GoogLeNet model trained on original-raw images, the GoogLeNet model trained on segmented images achieved 2.61% better performance for the "bir" dataset, 22.59% better performance for the "par_ozel" dataset, and 27.28% better performance for the "bir_ozel" dataset. Thus, one can conclude that the GoogLeNet model trained with segmented images achieves up to 27% higher performance. Furthermore, the GoogLeNet model was found to be more successful in the classification of colon tumor images as benign or malignant. On the other hand, in cases with non-homogenous data distribution between training and test datasets (par_ozel and bir_ozel datasets), the standard deep learning models were observed to perform with very poor accuracy, whereas deep learning models based on the proposed approach achieved very high levels of accuracy. In other words, deep learning models used for the classification of colon tumor images would perform much better if trained on images segmented with the augmented k-means clustering algorithm proposed in the first part of the study, instead of original-raw images.

Deep learning methods perform much better with larger datasets. Therefore, subsequent studies can utilize much larger datasets to build on the findings of the present analysis. In addition to the AlexNet and GoogLeNet models employed in the present study, further research can use other models such as ResNet or DenseNet. Moreover, the performance of the initial cluster center determination method employed in the augmented k-means clustering algorithm proposed for the segmentation of the images, was compared against those of random and non-linear methods. Further research can engage in comparisons with other cluster initialization methods such as random sampling, distance optimization, or density estimation. However, a fast and stable method with lower

processing requirements would be a better choice. For, the images to be segmented through the chosen method will be used as input for deep learning models. The processing time and processing load involved in the segmentation stage will have consequences for the deep learning model and the overall system.

5 Author contribution statements

In the scope of this study, the Author 1 in the formation of the idea, conceptualization, investigation, validation, visualization, writing, the literature review, and in the assessment of obtained results; the Author 2 supervision, checking the article in terms of content, and in the assessment of obtained results; the Author 3 checking the article in terms of content, in the assessment of obtained results, and supplying the microscope images were contributed.

6 Ethics committee approval and conflict of interest statement

This study was conducted using pathologic images, for which ethics committee approval was obtained with Decision No. 2015/116 of Necmettin Erbakan University, Non-drug, and Non-Medical Device Research Ethics Committee.

There is no conflict of interest with any person/institution in the article prepared.

7 References

[1] Türkyılmaz M, Hacıkamiloğlu E, Baran Deniz E, Boztaş G, Dündar S, Kavak Ergün A, Sevinç A, Tütüncü S, Seymen E, İlter H (Ed.), Keskinkılıç B(Ed.). "Türkiye Kanser İstatistikleri 2015". Halk Sağlığı Genel Müdürlüğü, Sağlık Bakanlığı, Ankara, 2018.

[2] Türkyılmaz M, Oruç Hamavioğlu Eİ, Dündar S, Kavak Ergün A, Sevinç A, Tütüncü S, Tolunay T (Ed.), Kaygusuz S (Ed.), Keskinkılıç B (Ed.), Özsarı S (Ed.), Gökler ME (Ed.). "Türkiye Kanser İstatistikleri 2018". Halk Sağlığı Genel Müdürlüğü, Sağlık Bakanlığı, Ankara, 2022.

[3] Boyle P, Levin B. "World cancer report 2008". International Agency for Research on Cancer, France, IARC Press, 2008.

[4] Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. "Global cancer statistics 2018: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries". *CA: a cancer journal for clinicians* 68(6), 394–424, 2018.

[5] Globocan. *Cancer incidence and mortality world-wide*. http://globocan.iarc.fr (06.06.2024)

[6] Yurtsever U, Evirgen H, Avunduk MC. "A new augmented k-means algorithm for seed segmentation in microscopic images of the colon cancer". *Tehnicki Vjesnik-Technical Gazette*, 25(2), 382–389, 2018.

[7] Isik H, Sezgin E, Avunduk MC. "A new software program for pathological data analysis". *Computers in biology and medicine*, 40(8), 715–722, 2010.

[8] Altunbay D, Cigir C, Sokmensuer C, Gunduz-Demir C. "Color graphs for automated cancer diagnosis and grading". *IEEE Transactions on Biomedical Engineering*, 57(3), 665–674, 2009.

[9] Yurtsever M, Yurtsever U. "Use of a convolutional neural network for the classification of microbeads in urban wastewater". *Chemosphere*, 216, 271–280, 2019.

[10] Parelanickal, S. B., & Jefferson, J. "Detection of Colorectal Cancer from Histopathological Images Tissue Classification Using Deep Learning Techniques". In 2023 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES) (pp. 1-7), 2023.

[11] Sari, M., Moussaoui, A., & Hadid, A. "Deep Learning Techniques for Colorectal Cancer Detection: Convolutional Neural Networks vs Vision Transformers". In 2024 2nd International Conference on Electrical Engineering and Automatic Control (ICEEAC) (pp. 1-6), 2024.

[12] Peng, C. C., & Lee, B. R. "Enhancing colorectal cancer histological image classification using transfer learning and ResNet50 CNN Model". In 2023 IEEE 5th Eurasia Conference on Biomedical Engineering, Healthcare and Sustainability (ECBIOS) (pp. 36-40), 2023.

[13] Anju, T. E., & Vimala, S. "Tissue and tumor Epithelium classification using fine-tuned deep CNN models". International Journal of Advanced Computer Science and Applications, 13(9), 2022.

[14] Babu, K. K., Kishan, K., Babu, B. S. V., Reddy, A., Chimma, P., Paruchuri, P. K. "Colon Cancer Nuclei Classification with Convolutional Neural Networks". Communications in Computer and Information Science, 376-385, 2024.

[15] Kumar, V. R. P., Arulselvi, M., & Sastry, K. B. S. "Comparative assessment of colon cancer classification using diverse deep learning approaches". Journal of Data Science and Intelligent Systems, 1(2), 128-135, 2023.

[16] Kumar, V. R. P., Arulselvi, M., & Sastry, K. B. S. "War strategy optimization-enabled Alex Net for classification of colon cancer". In 2022 1st International Conference on Computational Science and Technology (ICCST) (pp. 402-407), 2022.

[17] Cheng G, Guo W. "Rock images classification by using deep convolution neural network". *The 2nd Annual International Conference on Information System and Artificial Intelligence (ISAI2017),* China, 14–16 July 2017.

[18] Sachin R, Sowmya V, Govind D, Soman K. "Dependency of various color and intensity planes on cnn based image classification". *Third International symposium on signal processing and intelligent recognition systems*, Manipal, India, 13-16 Sep. 2017.

[19] Diaz-Cely J, Arce-Lopera C, Mena JC, Quintero L. "The effect of color channel representations on the transferability of convolutional neural networks". *Computer vision conference CVC 2019*, Las Vegas, 25-26 April 2019.

[20] Khojasteh P, Aliahmad B, Kumar DK. "A novel color space of fundus images for automatic exudates detection". *Biomedical Signal Processing and Control*, 49, 240–249, 2019.

[21] Yurtsever U. Detecting colon cancer using deep learning on segmented histopathological images. PhD Thesis, Sakarya University, Sakarya, Turkey, 2019.

[22] Ekicioğlu G, Özkan N, Salvaazar E. " Hematoksilen-eozin (hematoxylin-eosin)(h&e)". *Aegean Pathology Journal*, 2, 58–61, 2005.

[23] Lillie R, Fullmer H. *Histopathologic technic and practical histochemistry*. 3nd ed. 1976.

[24] Sirinukunwattana K, Pluim JP, Chen H, Qi X, Heng PA, Guo YB, Wang LY, Matuszewski BJ, Bruni E, Sanchez U, et al. "Gland segmentation in colon histology images: The glas challenge contest". *Medical image analysis*, 35, 489–502, 2017.

[25] Sirinukunwattana K, Snead DR, Rajpoot NM. "A stochastic polygons model for glandular structures in colon histology images". *IEEE transactions on medical imaging*, 34(11), 2366–2378, 2015.

[26] Nvidia-Digits. *The nvidia digits 6.0 software*. https://developer.nvidia.com/digits (06.08.2020).
[27] Nvidia-Cuda. *The nvidia cuda toolkit 9.0 software*. https://developer.nvidia.com/cuda-toolkit-archive (06.08.2020).
[28] Nvidia-cuDNN. *The nvidia cuda® deep neural network library (cudnn) software*. https://developer.nvidia.com/cudnn (06.08.2020).
[29] Jia Y, Shelhamer E, Donahue J, Karayev S, Long J, Girshick R, Guadarrama S, Darrell T. "Caffe: Convolutional architecture for fast feature embedding". arXiv preprint arXiv:1408.5093, 2014.
[30] Nvidia-Caffe. *The nvidia caffe software*. https://github.com/NVIDIA/caffe (06.08.2020).
[31] Albayrak S. "Color quantization by modified k-means algorithm". *Pakistan Journal of Applied Sciences*, 1(4), 508–511, 2001.
[32] Albayrak S, Karslıgil M. "The color clustering in color images with weighted k-means method". *9th Signal Processing and Application Congress SIU-2001*, Gazi Mağusa, KKTC, 25-27 Nisan 2001.
[33] Goodfellow I, Bengio Y, Courville A. *Derin Öğrenme*. Translated by Yarman Vural F, Cinbis RG, Kalkan S. Buzdağı Yayınevi, Ankara, Turkey, 2018.
[34] Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, Van Der Laak JA, Van Ginneken B, Sanchez CI. "A survey on deep learning in medical image analysis". *Medical image analysis*, 42, 60–88, 2017.
[35] Lopez AR, Giro-i Nieto X, Burdick J, Marques O. "Skin lesion classification from dermoscopic images using deep learning techniques". *13th IASTED International Conference on Biomedical Engineering (BioMed)*, Insbruck, Austria, 20-21 Feb. 2017.
[36] Bilginer O, Tunga B, Demirer RM. "Classification of skin lesions using convolutional neural networks". *Pamukkale University Journal of Engineering Sciences*, 28(2), 208-214, 2022.
[37] Akalın F, Yumusak N. "Classification of acute leukaemias with a hybrid use of feature selection algorithms and deep learning-based architectures". *Pamukkale University Journal of Engineering Sciences,* 29(3), 256-263, 2023.
[38] Kızrak MA, Bolat B. "A comprehensive survey of deep learning in crowd analysis". *International Journal of Informatics Technologies*, 11(3), 263–286, 2018.
[39] LeCun Y, Bottou L, Bengio Y, Haffner P, et al. "Gradient-based learning applied to document recognition". *Proceedings of the IEEE*, 86(11), 2278–2324, 1998.
[40] Krizhevsky A, Sutskever I, Hinton GE. "Imagenet classification with deep convolutional neural networks". *Advances in neural information processing systems*, pp. 1097–1105, 2012.
[41] Zeiler M, Fergus R. "Visualizing and understanding convolutional networks". *In European conference on computer vision*, Zurich, Switzerland, 6-12 Sep. 2014.
[42] Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A. "Going deeper with convolutions". *Proceedings of the IEEE conference on computer vision and pattern recognition*, Boston, MA, USA, 7-12 June 2015.
[43] Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A, et al. "Going deeper with convolutions". *arxiv 2014*. arXiv preprint arXiv:1409.4842, 1409, 2014.
[44] Simonyan K, Zisserman A. "Very deep convolutional networks for large-scale image recognition". *arXiv preprint* arXiv:1409.1556, 2014.
[45] He K, Zhang X, Ren S, Sun J. "Deep residual learning for image recognition". *IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* Las Vegas, NV, USA, 27-30 June 2016.